知识管理论坛 ISSN 2095-5472 CN11-6036/C
Knowledge Management Forum

E-mail: kmf@mail.las.ac.cn
http://www.kmf.ac.cn

【学术探索】

高校科研项目周期中数据质量控制模式探究

◎唐晶[1] 屈文建[2]
[1] 武汉大学信息管理学院 武汉 430061
[2] 南昌大学管理学院信息管理系 南昌 330031

摘要：[**目的/意义**]旨在为高校科研项目周期中数据质量控制提供一种有效控制途径与方法。[**方法/过程**]围绕科研项目周期和数据质量控制周期构建数据质量控制架构体系，并在此体系之下从认知、管理、流程3个视角分别实施数据质量控制，引入质量差距模型、企业架构模型和流程分析等控制方法，剖析高校科研项目周期中科研数据质量控制机理。[**结果/结论**]构建了适用于科研项目周期中高校科研数据质量控制架构体系及其数据质量控制模式，为高校科研数据质量控制提供理论支撑。

关键词：数据质量　科研项目周期　科研环境　质量差距模型　企业架构模型
分类号：G250

引用格式：唐晶，屈文建．高校科研项目周期中数据质量控制模式探究 [J/OL]. 知识管理论坛，2020, 5(1): 24-35[引用日期]. http://www.kmf.ac.cn/p/197/.

1 引言

随着第四科研范式环境即数据密集型科研环境的形成，科研数据量成指数型迅猛增长，数据表现形式多样，对于科研人员来说，其所需存储、处理的数据量惊人，数据来源和数据结构繁多复杂，为推动科研项目的进程带来了很大的阻碍[1]。且科研项目中充斥的数据资源来源复杂、存储结构多样，各个科研项目团队的数据共享技术兼容性差、数据管理意识不足，导致科研数据质量低下、利用率不高。而高校具有得天独厚的科研数据资源优势，是产生、利用科研数据从事科学研究的主力军，因此，高校更需要加强数据密集型科研环境下数据质量控制，并以此为基础探索、构建、完善国内高校在数字化科研背景下的科研数据管理架构与体系。

回顾国内外高校科研数据质量控制的研究

基金项目：本文系国家自然科学基金青年项目“基于学术异质网络表示学习的知识群落发现”（项目编号：71804135）和江西省社会科学“十三五”规划项目“高校跨学科团队知识共享特征与知识交流模式研究”（项目编号：18GL5）研究成果之一。
作者简介：唐晶 (ORCID：0000-0002-1211-5812)，硕士研究生，E-mail：1426137493@qq.com；屈文建 (ORCID：0000-0002-1043-4676)，教授，博士，硕士生导师。
收稿日期：2019-12-13　发表日期：2020-02-13　本文责任编辑：刘远颖

知识管理论坛，2020（1）:24-35
DOI：10.13266/j.issn.2095-5472.2020.003

历程，研究成果集中于理论体系的建立以及实践应用。在高校科研数据质量控制理论体系构建方面，DAMA 将数据质量控制分为 10 个职能域，成为当前较为被认可的数据管理知识体系 [1]；包冬梅等创新性地提出了高校图书馆数据治理框架——CALib 框架，促进了高校图书馆数据资源有效利用 [2]；屈文建等基于 Krantz 理论建立了高校科研数据质量控制架构并剖析其作用机制，为高校科研团队控制科研数据提供了有效理论支撑 [3]。在实践应用方面，马波等通过借鉴和总结欧美高校图书馆的科研数据管理实践经验，提出了加强数据治理意识与基础设施构建、人才培养等策略，以期提高国内高校科研数据治理水平 [4]；徐慧芳等通过调研国内科研数据管理实践并进行 SWOT 分析，从多个视角提出了国内高校图书馆有效实施科研数据服务与质量控制的途径 [5]。然而通过对已有研究的梳理可以发现，在数字化科研背景下，科研数据管理作为高校图书馆及科研机构新的业务生长点，其数据质量控制模式在应付海量的数据规模时还有较大的提升空间。

基于此，本文围绕科研项目周期构建高校科研数据质量控制架构体系，并从认知、管理、流程 3 个视角提出高校科研数据质量控制模式，为促进我国科研机构、图书馆等开展科研数据质量工程提供理论参考。

2 科研项目周期中实行数据质量控制的必要性

2.1 科研项目周期内涵及其相关数据活动

科研项目生命周期是一种管理学意义上的项目生命周期，描述科研项目“计划—立项—实施—结题”的发展阶段和历程。为便于后续研究，本文采用陆玉泉提出的科研项目生命周期划分方法，将科研项目周期划分为科研构思、科研实施、成果整理和成果发表 4 个阶段 [6]，如图 1 所示，基于此来研究科研项目周期中的高校科研数据质量控制。

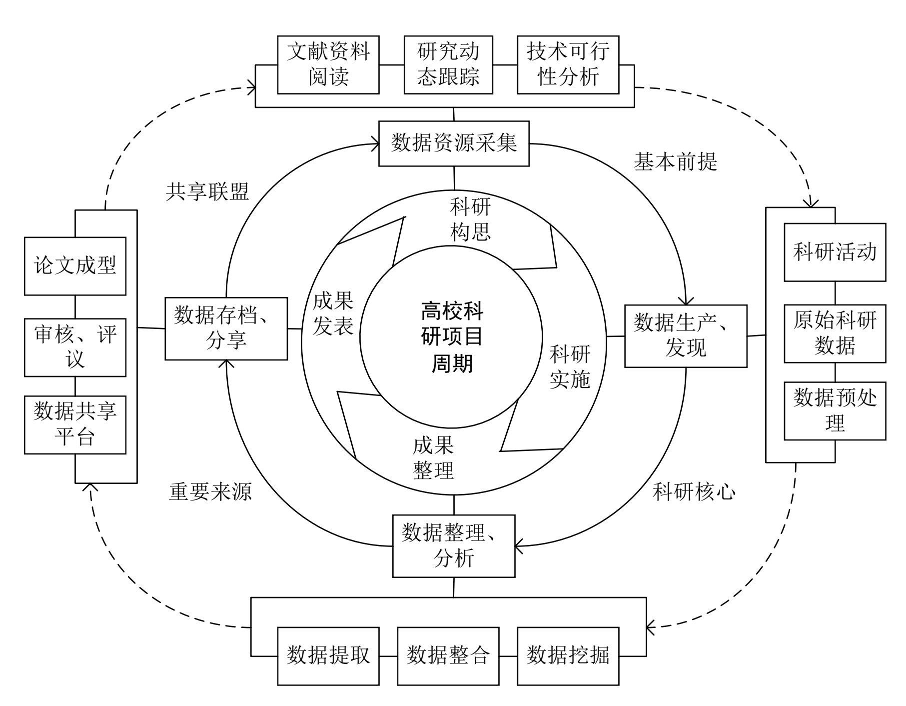


图 1 高校科研项目周期中相关数据活动

（1）科研构思阶段。科研构思阶段是科研项目选题和申报立项的基础，它体现了科研项目团队的研究能力和科研项目的新颖性、前沿性。这一阶段的数据活动主要是数据资源采集。

项目团队需要进行大量科研数据搜集、确定项目干系人、科研项目选题及数据质量规划等前期研究相关工作，其中数据搜集来源主要依靠共享联盟、主动上传和直接获取等方式[7]，采集内容包括但不限于：相关科研文献资料阅读、国内外发展方向与研究动态跟踪、技术可行性分析、市场调研等。

（2）科研实施阶段。在科研实施阶段，项目团队需要根据科研构思阶段完成的科研规划来进行学术论文撰写、科学实验开展等相关科研活动。这一阶段是科研数据生产、发现的黄金时期，需要按照前期行动中得到的数据准则严格规范新建数据，采取合适的数据处理工具对数据进行预处理，尽可能在早期获得较高质量的数据。

（3）成果整理阶段。得到原始科研数据后，需借助数据挖掘、数据库、数据质量检测识别等技术工具对数据进行提取、整合和挖掘，得到更精准的实验结果，并以此为根据撰写科研成果。对科研数据进行整合、分析与可视化，使其更直观地反映事物之间的真实联系是成果整理阶段的工作重心。

（4）成果发表阶段。经过前期的努力与锤炼，项目团队的科研成果成形，经修改、调整后进入评议、审查和发表阶段。这一阶段主要由项目评审小组对科研成果质量进行评议，审核通过后将科研数据上传至共享联盟等科研数据共享平台，进行数据的存档和分享，同时科研项目团队在经短暂休整后着手下一个科研项目的开展。

2.2 科研项目周期中高校科研数据质量问题及原因

科研数据作为高校科研项目周期中的重要组成，尽管其质量控制的重要性不断被学者强调，但在高校科研项目实践中仍存在诸多问题。数据错误、不完整、不一致、不及时是高校科研数据质量的主要问题（见表 1），其原因分为以下几种：①科研数据来源复杂且类型多样[6]。科研项目团队成员大多来自不同学科，拥有不同学科背景，由此而生的科研数据来源复杂且未能有统一数据规范，数据孤岛现象加剧，给数据的采集和整理带来不小挑战。②科研数据规模偏小。高校科研项目团队产生的科研数据规模不大，但项目团队数量众多，使得项目团队在数据发现中很难检索到高质量数据。③科研数据质量控制意识不足。项目团队成员的数据质量管理的习惯尚未养成，在科研数据搜集能力、数据共享意识、数据安全意识等方面还存在一定欠缺。

表 1 科研项目周期中主要数据质量问题

科研项目周期	主要数据活动	主要数据质量问题	
		原始数据	次级数据
科研构思	数据采集	数据采集不及时，丧失时效性； 数据记载存在遗漏情况； 数据采集格式不正确； ……	检索数据结果时效性差； 数据采集内容不符主题或错误； 检索数据结果不全面、不完整； ……
科研实施	数据生产、发现	实验数据记录与实际不符； 数据存储格式不规范； ……	数据录入错误或更新失败； 数据存储格式不规范； ……
成果整理	数据整理、分析	实验数据因更新不及时而呈现数据冗余； 实验数据分析错误； ……	数据报告内容不完整； 分布式数据内容不一致； 数据处理错误； ……
成果发表	数据存档、分享	共享数据内容不一致； 数据归档错误； ……	共享数据内容不一致； 数据归档错误； ……

知识管理论坛，2020（1）:24–35
DOI：10.13266/j.issn.2095–5472.2020.003

综合科研项目周期中数据质量问题剖析和原因分析可知，要将无序、信息混乱的科研数据序化整合以供高校科研项目团队重复利用，就必须针对存在的质量问题提出一套基于准确性、完整性、一致性和及时性的可靠数据质量评估标准，构建数据质量工程的实践架构体系，实现科研数据的创新增值、科研项目管理的效率提升和项目团队质量控制意识的提高。

2.3 实施高校科研数据质量控制的意义

结合高校科研数据自身特点、科研数据需求及科研环境等多方面情况，本文认为在高校科研项目周期中实施数据质量控制的意义主要表现在以下几个方面：

（1）促进跨学科科研数据的组织与融合。由于科研项目研究的需要和项目成员自身学科背景的异质性，科研数据普遍具有学科交叉和关联的特点，这一特征使得数据管理相关标准难以制定或缺乏普适性，造成科研数据利用率低下。在科研项目周期中实施数据质量控制有助于实施数据多维度组织，从不同维度对科研项目周期中不同阶段的科研数据进行组织与分类，提高科研数据利用率。

（2）满足高校科研项目团队对高质量数据的需求。对于高校科研项目团队人员而言，有缺陷的、不完整的或误导性的数据极有可能造成科学实验的失败，造成科研成本急剧上升，甚至项目无法结题。而准确、完整的科研数据在推动科研项目进程方面可起到事半功倍的效用，缓解项目团队的科研压力，提高数据的研究和使用价值。

（3）融入第四科研范式环境的必然要求。全新的科研范式下科研数据成为科学研究的基础性资源[8]。已有的科研数据是新研究的宝贵资产，对已有数据的整合、挖掘和再利用为学术研究提供了一种新的资源[9]。由此而生的高密集科研大数据对传统的科研数据管理造成了巨大冲击，因此在科研项目周期中构建高校科研数据质量控制架构体系、提高科研数据质量是融入第四科研范式环境的必然要求。

3 科研项目周期中高校科研数据质量控制架构体系

高校科研数据质量控制是一项时间跨度长、过程复杂的项目，在科研项目周期的不同阶段，科研数据的来源、质量要求和控制方法也不同。本研究从科研项目周期和质量控制阶段两个角度构建高校科研数据质量控制架构体系，如图2所示。高校科研项目周期中的相关数据活动涉及项目规划者、数据采集者等5类数据质量角色，他们在促进科研项目顺利开展的同时，从评估、认知和行动3个阶段开展数据质量控制活动，根据评估阶段产生的科研数据质量评估结果确定数据质量问题产生的根本原因，并从认知、管理、流程3个视角制定科研数据质量提高方案，从而形成科研数据质量控制周期，不断提高科研数据质量。

3.1 高校科研数据质量控制过程中相关角色

科研项目周期中主要涉及项目规划者、数据采集者、技术支持者、数据分析者和记录管理人员这5类角色，他们与科研项目和质量控制周期紧密相连，影响着科研数据的质量和价值。

（1）项目规划者。规划者一般包括科研项目中的项目主持人、数据架构师等人，主要在科研构思阶段收集科研项目数据需求，拟定科研项目数据管理规划，将数据质量提高融入项目设计和体系架构中，以便减少风险、抓住机遇和合理地分配资源。

（2）数据采集者。主要是指科研项目团队中实施科研实验或搜集数据的人员，他们掌握着数据的最初来源，作为数据质量角色的重要组成部分承担捕获、创建或维护数据的职责。

（3）技术支持者。主要指项目团队中数据库管理员、IT支持者等人员，开发科研数据存储平台，实施元数据管理并建立数据有效性规则，在逻辑上形成完整的资源体系，使科研数据在物理存储上具有整体性和关联性、在空间上具有较强的扩展性。



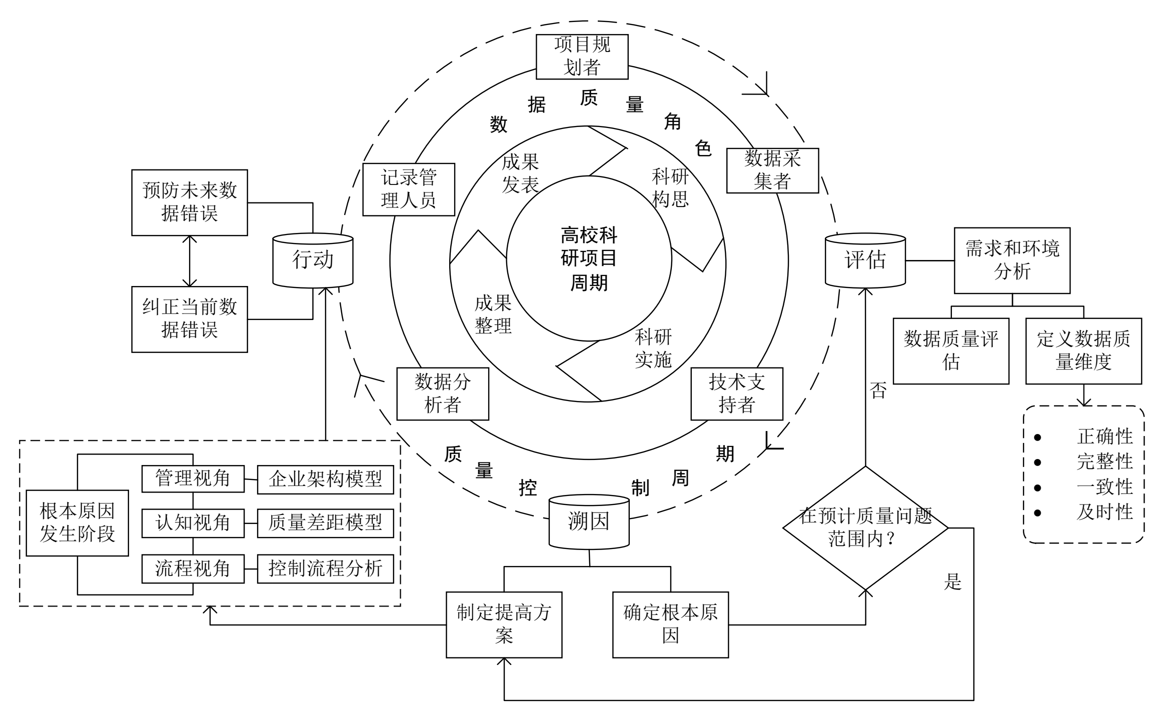


图 2 科研项目周期中数据质量控制架构体系

（4）数据分析者。知识工作者在成果整理阶段运用数据挖掘、数据建模等数据分析技术，对原始科研数据进行深入分析，识别、检测并清洗错误、无效等低质量数据，对高质量的数据进行整理、分析，挖掘科研数据背后事物之间的潜在联系，得出理想的科研成果。

（5）记录管理人员。记录管理人员在整个科研项目周期和质量控制周期中负责采集科研数据与规范存档记录，并明确其在生命周期各阶段的状态。由于数据通常分散在项目团队不同成员手中，因此建立科研数据在每个周期阶段的状态、操作、环境信息等完整的文档记录更为重要，以便减少科研数据在项目周期中的流失，加强了科研数据的管理与控制 [10]。

3.2 数据质量控制周期

数据质量控制周期通过评估、溯因和行动 3 个高层步骤，为讨论和着手思考“控制”提供一种简单的方法 [11]。在评估阶段，需要评判当前科研项目环境和数据，并将其与需求和预期相比较；在溯因阶段，了解数据和信息的真实状态、对科研的影响和根本原因是重点；前两个阶段的顺利进行为行动阶段打下坚实基础，数据质量管理人员在这一阶段除了纠正当前科研数据错误之外，还需预防未来信息和科研数据质量问题。评估是溯因的关键，溯因将引起行动，行动又由定期评估来核实，由此，科研数据质量控制得以良性循环。

3.2.1 评估阶段

评估阶段是整个质量控制周期的基础，也是高校科研项目周期得以持续的前提。在这一阶段主要包括科研数据的需求和环境分析、定义数据质量维度和数据质量评估 3 个方面。有效的规划是确保任何科研项目成功实施必不可少的要素，对科研数据的需求和环境进行全面分析将使科研项目实施真正做到有的放矢和更有针对性；根据科研项目周期中的数据质量问题并结合以往研究，定义科研数据质量维度包括正确性、完整性、一致性和及时性 4 个维度，统一规范科研项目周期中数据质量控制的标准和测量方法，以便后续数据质量评估的顺利开展。

3.2.2 溯因阶段

找出科研数据质量问题的根本原因并有针

知识管理论坛，2020（1）:24-35
DOI：10.13266/j.issn.2095-5472.2020.003

对性地制定提高方案是溯因阶段的重中之重。数据管理人员收集所有与低数据质量相关的背景信息，追踪和回溯科研数据在科研项目周期里流通的路径并进行捕获和比较，对质量问题进行优先级分类，最终确定数据质量问题的根本原因。若根本原因引起的质量问题超出了预先的数据质量评估维度，则需返回评估阶段重新对数据需求和环境进行分析；若根本原因在预计质量问题范围内，需要确定根本原因发生的阶段，提出解决根本原因的具体建议，并对若干建议划分优先等级，确定质量提高方案。

3.2.3 行动阶段

行动阶段的目的是纠正当前科研数据错误，同时采取手段预防未来科研数据错误。由溯因阶段输入至行动阶段的质量提高方案主要包括质量差距模型、企业架构模型和流程分析 3 种模式。从认知视角来看，引入质量差距模型，认知并分析期望的科研数据质量和实际感知的科研数据质量之间的差距，对于正确把握质量问题产生的根本原因具有关键作用；从管理视角来看，随着跨学科趋势的增强，高校科研项目更强调跨部门性的协同合作，在企业架构的模型上实施高校科研数据质量控制能从科研团队层面为数据质量控制活动提供保障；从流程视角来看，高校科研数据质量水平还与项目周期中各个环节息息相关，因此本研究还引入流程分析，从微观角度有计划、有组织地把控高校科研数据质量。这 3 个视角提出的科研数据质量改进流程和数据纠正方法，为科研项目周期中高校科研数据质量控制提供了理论思路和实践基础。

4 科研项目周期中高校科研数据质量控制模式分析

为确保数据管理方可以自始至终控制科研项目周期中的科研数据质量，在产生实质影响前尽早识别、排序和修复数据缺陷，本文基于高校科研数据质量控制架构体系的三大阶段，提出了科研项目周期中高校科研数据质量控制总模式，如图 3 所示。该模式阐述了在高校科研数据质量控制架构体系下数据质量控制的具体模式及流程。评估阶段和溯因阶段是在全项目周期中实时监测科研活动中的数据质量，依据科研数据质量规则及维度对数据质量进行评估后，找出数据质量问题并根据问题的性质、出现的阶段进行分类编号。在行动阶段，数据管理人员将从认知视角出发，引入质量差距分析模型，将实际的数据质量与项目团队期望的数据质量进行比较，进行质量差距分析，从而深入测量、严密评判各类数据质量问题并找到根本原因，为科研团队制定数据质量控制决策提供关键依据；从管理视角来看，数据管理人员根据基于质量差距模型的科研数据质量控制模式排查出来的质量问题根本原因，将数据质量控制整合为企业架构的固有组件，从宏观上提出解决科研项目中跨学科、跨部门的数据质量控制和协同监管问题的方法，为科研项目周期中的高校科研数据质量控制活动提供理论支撑和数据治理规范；最后，数据管理人员从流程视角上对整个科研项目周期进行数据质量问题监测与跟踪，从微观层面确保问题能得到及时有效的处置，并提高项目流程管理效率。现对各视角下数据质量控制模式进行剖析。

4.1 基于质量差距模型的高校科研数据质量控制模式

营销学家 A. Parasuraman 和 L. Berry 等人提出服务质量差距模型，通过对顾客感知的实际服务质量与期望服务质量的差距比较来分析存在的质量问题 [12]，J. R. Evans 提出“客户驱动质量”方法，主张从需求的角度提升数据质量 [13]。在高校科研项目周期中，对科研数据管理人员和项目团队进行数据质量差距分析是发现数据质量根本原因、实现数据质量有效提高的关键前提。质量差距分析模型主要是通过项目团队期望的数据质量和其实际感知的数据质量的差距，进行根本原因分析，从而精准修复数据质量问题，实现提高科研数据质量的目的，具体如图 4 所示：

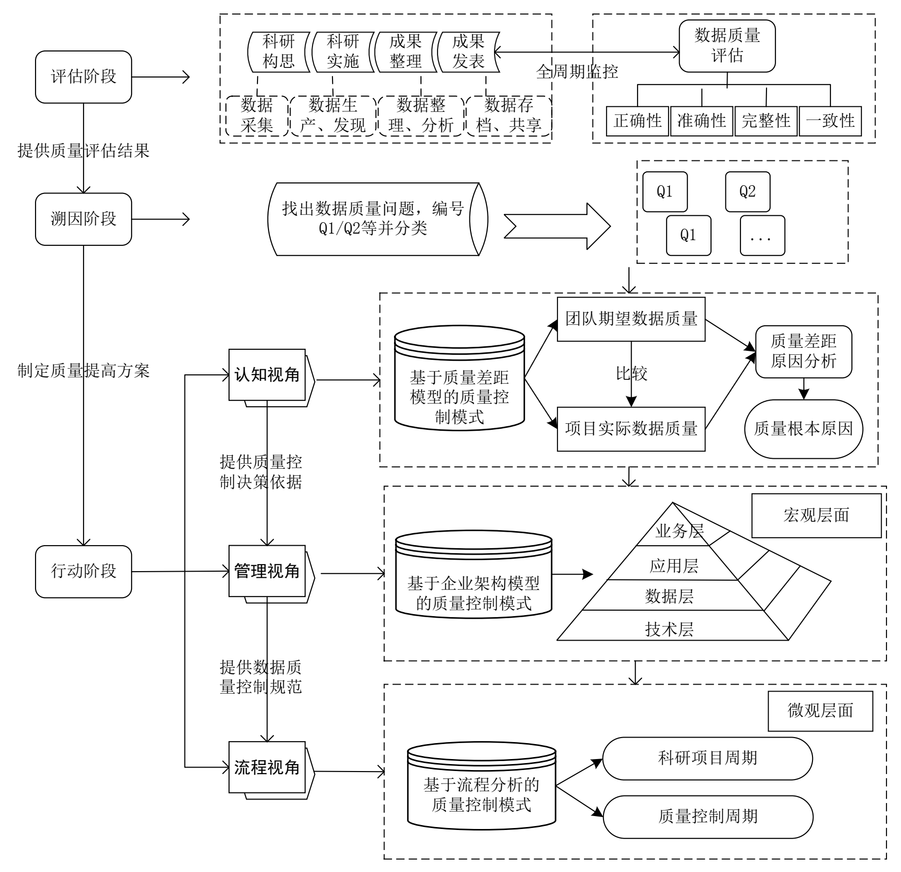


图 3　高校科研项目周期数据质量控制总模式

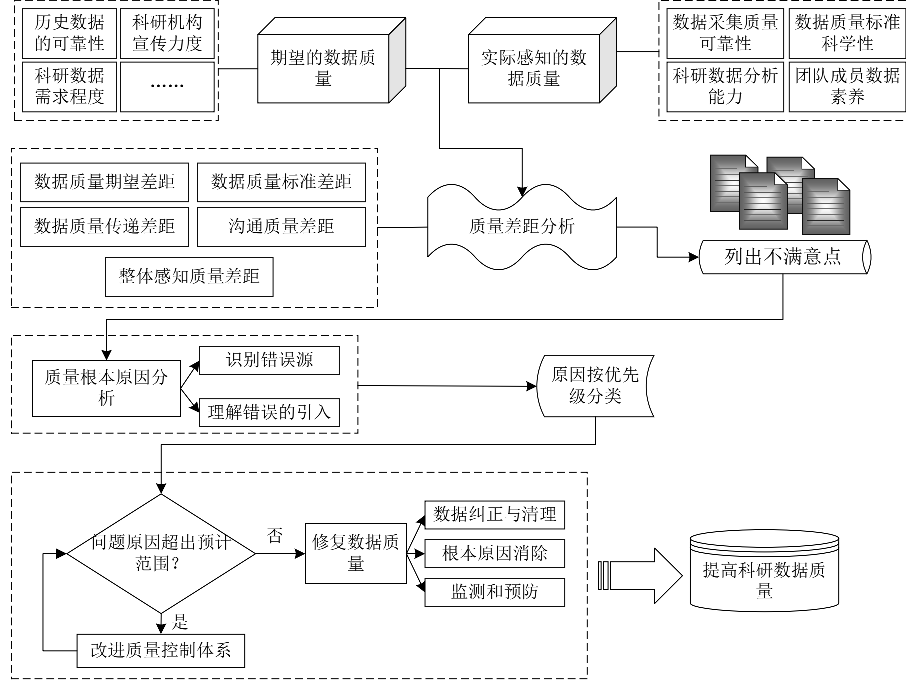


图 4　基于质量差距模型的质量控制模式

知识管理论坛，2020（1）:24–35
DOI：10.13266/j.issn.2095–5472.2020.003

影响项目团队成员期望的数据质量因素一般包括历史数据的可靠性、科研机构的宣传力度以及科研数据的需求程度等，这些因素的相互作用构成了科研项目周期中项目团队对科研数据质量的期望指数；而项目团队实际感知到的数据质量的形成来源于数据采集质量的可靠性、数据质量标准制定的科学性、对科研数据的分析能力和团队成员的数据素养。数据质量控制人员通过用户调查法、问卷调查法多种方式了解这两者之间的差距，并根据调查结果进行质量差距分析。质量差距分析内容由数据质量期望差距、数据质量标准差距、数据质量传递差距、沟通质量差距和整体感知质量差距构成，其中整体感知质量差距是整个模型分析的关键，前 4 个差距的大小直接影响着科研项目团队感知的科研数据质量差距的大小，因而在进行质量控制时应尽量缩小前 4 个差距。数据控制人员通过质量差距分析列出来的差距清单，进行快速配置测试或查询，记录下可能需要纠正的数据集，进行质量根本原因分析。识别错误源和理解错误的引入是根本原因分析的重点，主要通过在整个科研项目周期中重复识别数据"进时有效、出时无效"的处理阶段，精准定位数据活动中数据错误源的位置，在此基础上更进一步分析数据质量缺陷造成的原因，了解引入错误的方式。将数据质量原因按优先级分类，有助于将业务价值与高质量数据关联起来，同时还能发现少量在评估阶段疏漏的质量问题，促进数据质量控制体系的改进。若问题原因在预计范围内，则需要采取数据纠正与清理、根本原因消除、监测和预防等修复质量方法，以提高科研数据质量。

4.2 基于企业架构模型的高校科研数据质量控制模式

高校科研项目团队成员大多来自于不同院系，学科背景的异质性使得科研数据内容综合性强、跨学科性强，数据难以分类、有序储存；同时高校科研数据质量控制需要多个部门共同协作，具有跨部门性，因此从管理角度来说还需要构建一套适应高校科研数据特征的数据质量控制模式。企业架构是一种按照一定系统设计原则进行组织的理论框架，在信息资源集成 [14]、数据治理 [15] 等方面得到广泛应用，因此可将企业架构模型引入高校科研数据质量控制模式。企业架构的内容一般包括业务架构、应用架构、数据架构和技术架构 4 层 [16]，将其应用于科研数据质量工程实践中，具体如图 5 所示。

技术层是整个企业架构模式的基础，涉及到网络传输基本服务、IT 基础设施设备、大数据处理技术和智力支持等要素。其中网络传输和 IT 基础设施设备为支撑整个科研数据活动和质量控制活动提供了保障，大数据处理技术是数据密集型科研环境下获取高质量数据不可或缺的支撑点，为数据质量控制提供强大的技术支持，包括数据关联分析、用户行为智能预测、知识图谱分析、聚类分析等，可实现对原始科研数据的有效处理。

数据层具体可分为数据采集、数据质量控制和数据存储 3 个方面，该层面主要是为科研项目提供多维化的科研数据来源，并运用数据处理工具和技术对数据进行预处理和规范化存储。数据采集方面，凡是能满足项目团队科研需要的资源都成为数据采集的对象，这些多维化数据来源主要包括共享联盟资源、网络资源和项目历史资源，项目团队根据科研需要从中提取知识，并将其应用到实际科研活动之中。但鉴于数据质量的良莠不齐，应当按照数据质量控制周期的循环步骤对采集的数据进行质量控制，经评估后确定数据质量根本原因、确定质量提高方案，从而预防和纠正数据错误，完成数据预处理工作。同时，还应设置必要的数据存储规范，对数据进行统一格式化后再存档至科研数据存储平台。

应用层主要是根据业务层和数据层提供的数据质量控制方案进一步提高数据质量，即对科研数据进行提取、整合和深入挖掘。其中数据提取是指对科研数据进行选择和清洗，保证原始数据的准确性和及时性；数据整合是指对



知识管理论坛
2020 年 第 1 期(总第 25 期)

数据进行过滤和合并，降低数据的冗余性；数据挖掘是通过挖掘模式和算法应用对科研数据下的潜在知识内容进行挖掘和分析，最终得到理想的科研数据成果。

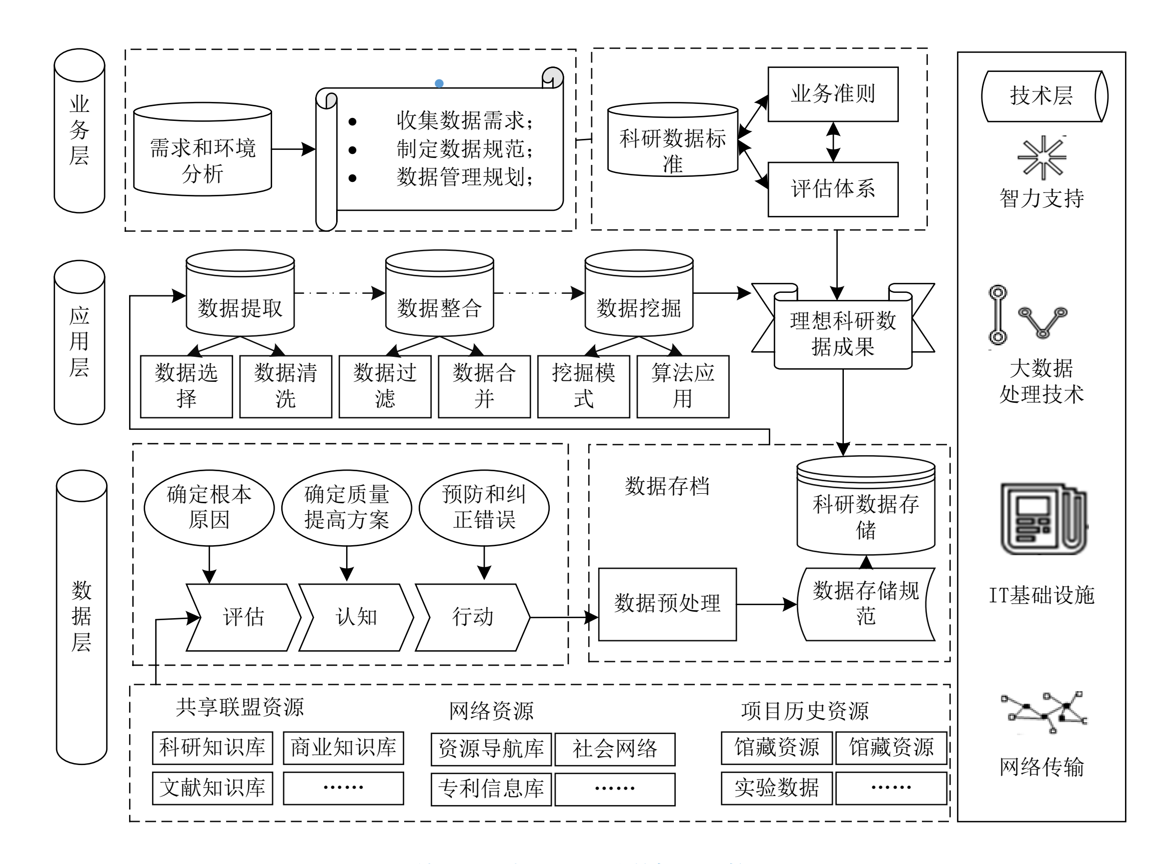


图 5 基于企业架构模型的数据质量控制模式

业务层则是为提高科研项目周期中的数据质量而形成的集成视窗，对整个项目中的数据活动进行实时监测和维护。除此之外，还需在科研构思阶段进行需求和环境分析，如收集数据需求、制定数据规范和数据管理规划等，同时制定适用于整个架构的科研数据标准，为科研数据质量的评估、控制提供有效且统一的尺度。

4.3 基于流程分析的高校数据质量控制模式

科研数据质量包括过程数据质量和结果数据质量，结果质量的好坏反映在科研项目过程的各个环节上，只有改善过程数据质量才能促使结果质量的提高。因此，需要对科研项目周期中的高校科研数据质量控制进行流程分析，从微观的角度有计划、有组织、持续地进行质量改进，达到期望数据质量的要求，如图 6 所示。

在科研构思阶段，数据资源采集、需求环境分析和质量评估标准的确定是规划者和记录管理人员的重点工作。数据资源的采集来源于科研数据共享平台，采集过程中从源头上保证数据的准确性与完整性是数据质量控制的重点；需求环境分析中主要是了解项目团队的科研数据质量需求、前沿性分析等，此时科研数据的准确性和及时性是规划者的主要需求；在制定数据质量评估标准时，规划者和记录管理人员需要充分考虑数据质量需求，确保标准的可靠性和科学性。

知识管理论坛，2020（1）:24–35
DOI：10.13266/j.issn.2095–5472.2020.003

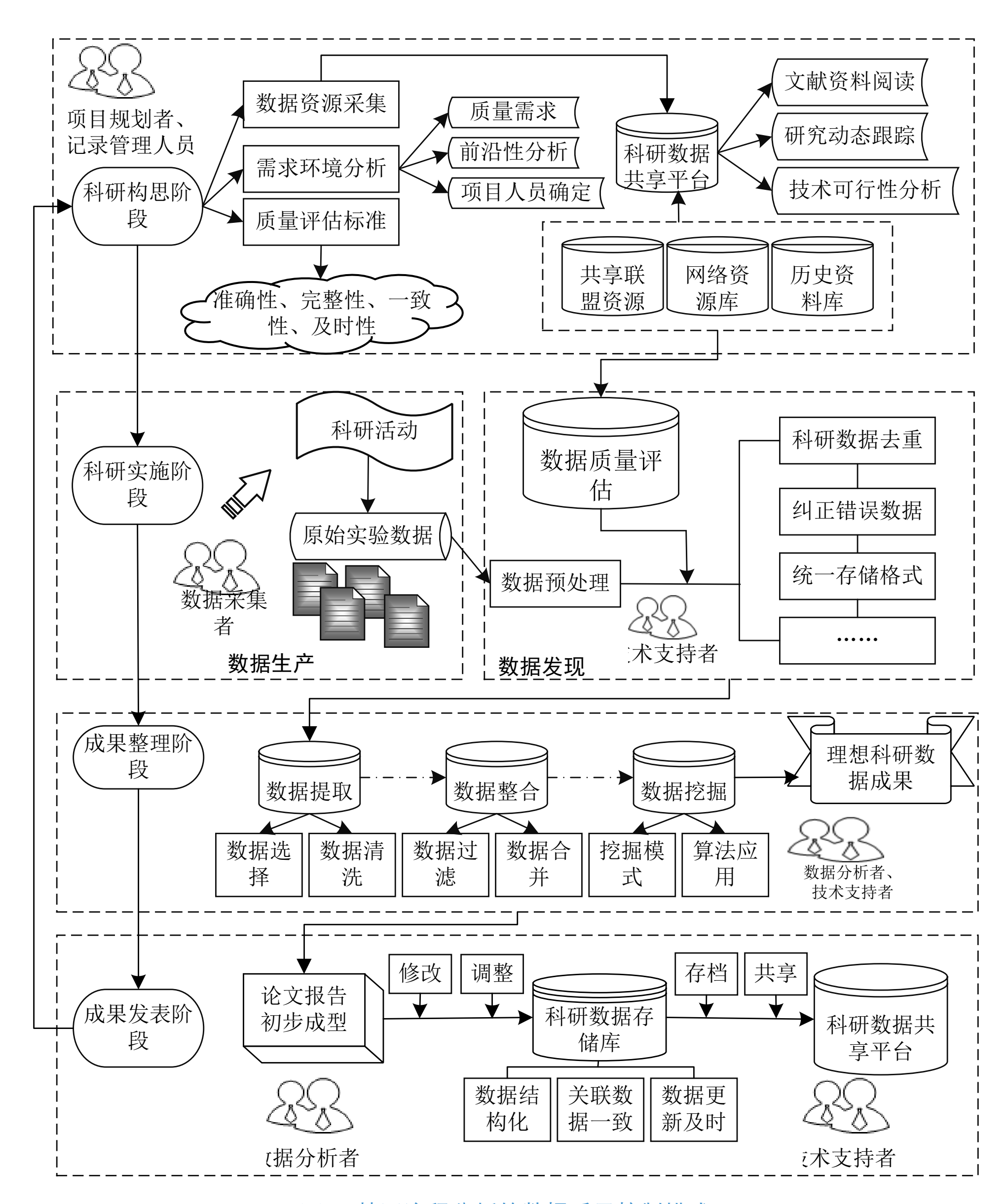


图 6 基于流程分析的数据质量控制模式

在科研实施阶段，生产者通过科研活动产生原始实验数据，由技术支持者根据数据质量评估结果，运用大数据处理技术进行数据预处理，包括科研数据去重、纠正错误数据、统一存储格式等内容，得到符合科研数据标准规范的结构化数据。这一阶段中严谨科学的科研活动是影响后续数据质量的根本要素，数据预处理的正确操作是提高科研数据质量的基本前提。

在成果整理阶段，从数据发现上升到知识发现、从数据采集到数据创造是这一阶段的主旋律。数据的提取、整合和挖掘依赖于技术支持者对数据处理技术的娴熟掌控和知识工作者的数据素养，两者的协同作用才能得到更精准的科研实验结果，才能更深入地揭示事物间潜在关联和规律。

在成果发表阶段，知识工作者对成果整

理阶段的科研数据进行知识整合与挖掘，得到论文报告基本雏形。剔除部分偏离正常值的数据，对科研成果进行适度修改调整后存入科研数据存储库中。科研数据存储库将数据结构化，保证关联数据的一致性，并实时更新数据，将高质量的科研成果存档、共享至科研数据共享平台。

5 结论

大数据时代下，科学研究第四范式归根结底是对海量数据的挖掘与控制 [6]，科研数据的质量控制也就成了科研人员根据科研数据进行深入研究的关键所在。本研究从科研项目周期内涵及相关数据活动阐述了引入数据质量控制的必要性，并从科研项目周期和质量控制阶段两个角度构建高校科研数据质量控制架构体系，在深入剖析体系中各部分的功能和特点基础上，从认知、管理和流程三大视角构建了各有侧重又综合协同的数据质量控制模式及流程，为高校科研项目周期中的科研数据质量控制提供了理论支撑和有效途径。

参考文献：

[1] DAMA 国际 .DAMA 数据管理知识体系指南 [M]. 北京 : 清华大学出版社 , 2012:4-9.

[2] 包冬梅 , 范颖捷 , 李鸣 . 高校图书馆数据治理及其框架 [J]. 图书情报工作 , 2015, 59(18):134-141.

[3] 屈文建 , 唐晶 , 陈旦芝 . 高校科研数据质量控制架构与机制研究 [J]. 情报理论与实践 , 2018, 41(11):45-50.

[4] 马波 , 李宇 . 欧美国家高校图书馆科研数据管理实践及启示 [J]. 图书馆工作与研究 , 2018(8):17-24.

[5] 徐慧芳 . 中国科研数据管理实践现状调研及对图书馆的启示 [J]. 图书馆学研究 , 2019(17):45-53.

[6] 陆玉泉 . 大数据时代高校科研数据服务模式与建设路径研究 [D]. 南京 : 南京航空航天大学 , 2017.

[7] 熊文龙 , 李瑞婻 . 基于科学数据管理的图书馆数据服务研究 [J]. 图书情报工作 , 2014, 58(22):48-53.

[8] 孙继周 .E - Science 环境下高校图书馆开展科学数据管理与共享的路径研究 [J]. 图书馆 , 2016(5): 66-71.

[9] MARCHIONINI G, 杨冠灿 , 芦昆 . 科研数据管理 : 保障数据质量 , 促进 iSchools 新科学研究 [J]. 图书情报知识 , 2013(4): 4-9.

[10] 丁宁 , 马浩琴 . 国外高校科学数据生命周期管理模型比较研究及借鉴 [J]. 图书情报工作 , 2013, 57(6):18-22.

[11] MCGILVRAY D. 数据质量工程实践 [M]. 曹建军 , 刁兴春 , 张健美 , 译 . 北京 : 电子工业出版社 , 2010.

[12] MBA 智库百科 . 诺伯特维纳的控制 [EB/OL]. [2012-04-14]. http//wiki.mbalib.com/wiki/Control_theory.

[13] EVANS J R, LINDSAY W M. The management and control of quality[M]. 5th ed.St.Paul:West, 2002.

[14] 屈文建 , 周小渝 , 孟书荣 . 基于企业架构的数字图书馆科技信息资源集成服务研究 [J]. 图书馆学研究 , 2017(13):32-39.

[15] 郭路生 , 刘春年 . 大数据时代应急数据质量治理研究 [J]. 情报理论与实践 , 2016, 39(11):101-105.

[16] CIOCouncil. Federal Enterprise Architecture Framework [EB/OL]. [2016-02-03]. http://www.enterprise-architecture.info/Images/Documents/Federal%20EA%20Framework.pdf.

作者贡献说明：

唐　晶：提出研究思路，设计研究方案，撰写论文；

屈文建：提出研究思路，指导论文修改。

知识管理论坛，2020（1）:24–35
DOI：10.13266/j.issn.2095–5472.2020.003

Research on Data Quality Control Mode in University Scientific Research Project Cycle

Tang Jin[1] Qu Wenjian[2]
[1]School of Information Management, Wuhan University, Wuhan 430061
[2]Department of Information Management, School of Management, Nanchang University, Nanchang 330031

Abstract: [Purpose/significance] This paper aims to provide an effective control approach and method for the data quality control in the scientific research project cycle of universities. **[Method/process]** It built a data quality and quality control architecture system around the scientific research project cycle and the data quality control cycle, and implemented data quality control from the perspectives of cognition, management, and process under this system, and introduced quality gap models and companies. Control methods such as architecture model and process analysis analyzed the data quality control mechanism in the scientific research project cycle of universities. **[Result/conclusion]** A scientific research data quality control architecture system and a data quality control model suitable for the scientific research project cycle were established, which provided theoretical support for the quality control of scientific research data in universities.

Keywords: data quality research project cycle research environment quality gap model enterprise architecture model